Cell, Vol. 84, 331–334, February 9, 1996, Copyright ©1996 by Cell Press

STATs: Signal Transducers and Activators of Transcription

Minireview

James N. Ihle
St. Jude Children's Research Hospital
Department of Biochemistry
Memphis, Tennessee 38105

Signal transducers and activators of transcription (STATs) were first identified as a unique family of DNA-binding proteins approximately four years ago. Since that time, the number of mammalian family members has grown, and in this issue of *Cell*, two reports now extend the STAT family to Drosophila (Hou et al., 1996; Yan et al., 1996). The STATs have drawn considerable attention because of their unique mode of activation and the diversity of biological effects they are thought to mediate from antiviral responses to cell transformation. As in any emerging field, there has been a rapid explosion of information and speculation. Now the dust is settling, and a more enduring picture is beginning to emerge as exemplified by two articles that deal with STAT1 function in this issue of *Cell* (Durbin et al., 1996; Meraz et al., 1996). This review provides a brief overview of STAT structure, function, and possible evolution. Several reviews (Darnell et al., 1994; Schindler and Darnell, 1995; Ihle, 1995; Ihle and Kerr, 1995; Ihle et al., 1995) provide more detail and, more importantly, references to primary information.

STAT Family Members

The first STAT family members were identified as DNA-binding proteins in interferon (IFN)-regulated gene expression. From these studies, two STATs and two distinct models of STAT involvement emerged. In response to IFNα/IFNβ, a DNA-binding complex is rapidly formed consisting of STAT1, STAT2, and a DNA-binding protein termed p48, which binds an IFN-stimulated response element (ISRE). In contrast, in response to IFNγ, a DNA-binding complex is formed of STAT1 homodimers that binds a unique element termed the IFNγ-activated sequence (GAS). Formation of either complex is dependent upon tyrosine phosphorylation of STAT2 and STAT1, or just STAT1, respectively.

Following the cloning of *Stat1* and *Stat2*, it became obvious that STAT-like activities were activated by various cytokines. This prompted efforts that resulted in the identification of five additional mammalian *Stat* genes (Figure 1). *Stat3* was cloned as an interleukin-6 (IL-6)-activated transcription factor or by homology to *Stat1*. *Stat4* was cloned by homology approaches. *Stat5* was cloned as a prolactin-activated transcription factor from sheep. Subsequently, it was found that in mice there are two highly related *Stat5* genes, *Stat5a* and *Stat5b*. Most recently, *Stat6* was cloned as an IL-4-activated DNA binding as well as by homology. Although well characterized in mammals, the existence of nonmammalian STATs has only recently been shown as reported by two groups in this issue of *Cell* (Yan et al., 1996; Hou et al., 1996).

Genetic mapping of the mammalian *Stats* (Copeland et al., 1995) suggests a potentially interesting evolution that may relate to their functions (detailed below). Specifically, the *Stats* colocalize with *Stat1–Stat4*, *Stat2–Stat6*, and *Stat3–Stat5a/Stat5b* tightly linked on chromosomes 1, 10, and 11, respectively. This suggests that there existed an evolutionarily primordial gene that duplicated and that this duplication, or the original, was further duplicated. In recent evolution, the *Stat5* gene further duplicated. Since the Drosophila gene appears to be most related to *Stat5*, the *Stat3–Stat5a/Stat5b* site may represent the ancestral gene site.

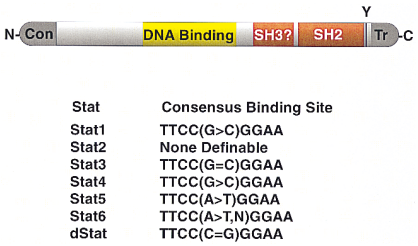
STAT: Structure and Functional Domains

The STATs share several conserved structural and functional domains (Figure 1). The most interesting and conserved domain is a potential phosphotyrosine-binding, SRC homology 2 domain (SH2). Several observations support this prediction, including the finding that mutation of the predicted phosphotyrosine-binding Arg residue eliminates activity. However, the isolated SH2 domains have yet to be shown to bind phosphotyrosine-containing peptides selectively. Nevertheless, the SH2 domain plays three important roles. It is critical for the recruitment of STATs to activated receptor complexes (detailed below). It is required for the interaction with the Janus protein-tyrosine kinases (JAKs), which phosphorylate the STATs. Finally, the SH2 domain is required for STAT dimerization and the associated ability to bind DNA. The STATs were also reported to contain SH3 domains. This region is much less conserved, including the critical residues involved in proline binding, and no evidence has emerged to suggest an SH3 function.

DNA binding by purified STAT1 is totally dependent upon tyrosine phosphorylation at a single site (Tyr-701), carboxyl to the SH2 domain. Similarly, the DNA-binding activity of all STATs is dependent upon tyrosine phosphorylation and, where examined, involves a comparably located tyrosine. Considerable evidence supports the hypothesis that tyrosine phosphorylation results in dimerization of STATs through the intermolecular interaction of the SH2 domains and the carboxyl sites of tyrosine phosphorylation and that this dimerization is essential for DNA binding.

The DNA-binding domain of STATs is located in the middle in a highly conserved domain that is not found in other DNA-binding proteins. Consistent with both domain conservation and the requirement for dimerization, all but one of the STATs bind very similar symmetrical, dyad sequences (Figure 1). Indeed, the consensus sequences, defined by using STAT homodimers in binding and amplification reactions with random oligonucleotides, only differ in the center core nucleotide or, in the case of STAT6, the preference for two central core nucleotides. Curiously, phosphorylated STAT2 does not, or weakly, binds DNA, suggesting that it may only function as a heterodimer with STAT1, in complex with the p48 DNA-binding component.

STATs may also form heterodimers. In addition to STAT1–STAT2 heterodimers, STAT1–STAT3 heterodimers are frequently observed in cytokine responses. In contrast, STAT4, STAT5, and STAT6 have yet to be shown to form heterodimeric complexes, although the



Stat	Consensus Binding Site
Stat1	TTCC(G>C)GGAA
Stat2	None Definable
Stat3	TTCC(G=C)GGAA
Stat4	TTCC(G>C)GGAA
Stat5	TTCC(A>T)GGAA
Stat6	TTCC(A>T,N)GGAA
dStat	TTCC(C=G)GGAA

Figure 1. Structure and Consensus DNA-Binding Sequences of STATs

The functional domains of the STATs are indicated, including a conserved region in the amino terminus (Con), the DNA-binding domain, a SRC homology domain 3–like region (SH3?), the highly conserved SRC homology domain 2 region (SH2), the critical site of tyrosine phosphorylation (Y), and the carboxy-terminal transcriptional activation domain (Tr). The consensus DNA-binding sites for the mammalian STATs (STAT1–STAT6) and the D-STAT (dStat) are indicated. These were determined in various studies by binding and amplification reactions using constructs containing a random sequence core. In each case, the sequences are those obtained with homodimers of the indicated STATs.

highly related STAT5a and STAT5b proteins form heterodimers (Quelle et al., 1996).

The highly divergent carboxy-terminal domain of the STATs is required for, or influences, transcriptional activation. In the case of *Stat1*, there exists a naturally occurring splice variant (STAT1β) that lacks the carboxyl 38 amino acids. This variant is recruited to the receptor complex, becomes phosphorylated and binds DNA, but does not activate gene transcription. Indeed, Stat1β acts as a naturally occurring dominant negative. Similar variants have been identified for STAT3 and STAT5. The divergent nature of the carboxyl region is likely to be critical, since the STATs all bind very similar DNA sequences but affect individual gene expression in unique ways. Thus, it can be hypothesized that unique transcriptional activation domains provide the specificity for function within the context of specific transcription complexes.

The functions of STATs may also be influenced by serine phosphorylation. The DNA-binding activity of STAT3 was shown to be affected (Zhang et al., 1995), and it was proposed to involve phosphorylation of a mitogen-activated protein kinase (MAPK) site in the carboxyl region. However, it is unlikely that a carboxyl site of phosphorylation would affect DNA binding; of more concern is the lack of confirmatory data of the initial observation. STAT1 has been shown to be phosphorylated at Ser-727, a potential MAPK site, and this phosphorylation influences transcriptional activation (Wen et al., 1995). Although this suggests a link between the RAS pathway and STATs, IFNs do not activate the RAS pathway. This conundrum might be resolved by the surprising observations that the MAPK Erk-2 directly associates with the IFN receptor, is activated by ligand binding, and can be directly immunoprecipitated with STAT1 (David et al., 1995). Clearly, these observations provide a series of firsts that must be confirmed.

The amino-terminal region of STATs is also conserved and critical to STAT function, since even small deletions completely eliminate the ability of STATs to be phosphorylated. Exactly how these domains contribute to STAT structure is not known. Finally, the domains, or mechanisms, required for STAT translocation to the nucleus are unknown.

STATs and the Cytokine Receptor Superfamily

One or more STATs are activated in response to all cytokines that utilize cytokine receptor superfamily members. This receptor family transduces signals for approximately 30 cytokines having diverse biological functions. They share conserved extracellular, ligand-binding motifs suggesting a common evolution. The family is also characterized by association with one or more members of the JAK family. Considerable evidence supports the hypothesis that ligand binding results in aggregation of the receptor chain(s) and the associated JAKs, allowing transphosphorylation and activation of JAK catalytic activity. The JAKs subsequently phosphorylate the receptors on tyrosines, as well as a variety of cellular substrates that are recruited to the activated receptor complexes. Consistent with this, JAK activation is required for initiation of multiple signaling pathways in response to cytokines including the RAS pathway in many cases.

The consistent activation of JAKs and STATs by cytokines has given rise to the concept of a JAK–STAT signaling pathway. This is somewhat of a misnomer since the JAKs are required for the activation of all signaling pathways, independent of the STATs. Nevertheless, the concept emphasizes the consistent association of JAK and STAT activation by cytokine receptors. Indeed, as detailed in this issue of *Cell*, this prompted genetics studies in Drosophila to identify a comparable pathway in flies based on the existence of a Drosophila homolog of the mammalian JAKs, the *hopscotch* gene. It now remains to be determined whether any homologs of the mammalian cytokine receptor superfamily, and their associated cytokines, exist in Drosophila. If a Drosophila cytokine receptor is involved, it would support the concept that this family of receptors has evolved with JAKs and STATs as a unique signaling system.

Specificity of Recruitment and Activation of STATs in Response to Cytokines

In the response to cytokines, there is often a remarkable specificity in STAT activation. For example, in lymphocytes, IL-2 activates STAT5, IL-12 activates STAT4, and IL-4 activates STAT6. Specificity is not controlled by the JAKs, but rather by the ability of individual, activated receptor complexes to recruit specific STATs to the receptor complex. This occurs through the interaction of the STAT SH2 domain with specific sites of receptor tyrosine phosphorylation (Figure 2). This has been elegantly demonstrated by the observation (Stahl et al., 1995) that short peptides, containing a “docking” site for STAT3, when added to a receptor will allow the recruitment and activation of STAT3. Conversely, simply swapping STAT SH2 domains can change the receptors to which the chimeric STATs are recruited and activated (Heim et al., 1995).

The above studies, and the many others that have followed, have given rise to a general model (Figure 2) in which the STAT SH2 domain initially recruits the STAT to the receptor complex at specific sites of tyrosine phosphorylation. The STATs are thus accessible to the activated JAKs and subsequently associate with the JAK and become phosphorylated. Dimerization is hypothesized to allow release of the STATs from the complex, and unknown mechanisms mediate nuclear trans-

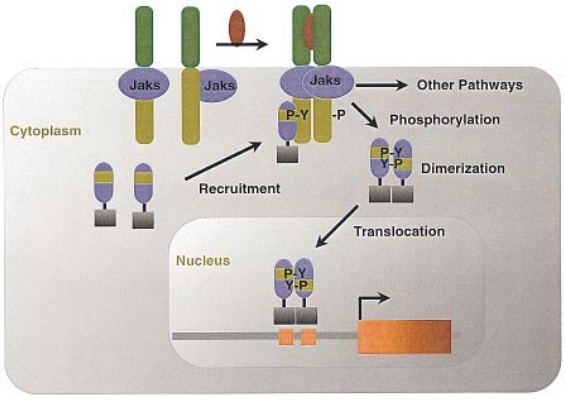


Figure 2. Mechanisms of Activation of STATs in Signaling by the Cytokine Receptor Superfamily

Ligand-induced receptor aggregation initiates the response by bringing the associated JAKs into sufficient proximity to allow trans-phosphorylation and activation of catalytic activity. The activated JAKs subsequently phosphorylate the receptors chains at multiple sites. The STATs are then recruited to the activated receptor complex through the interaction of the SH2 domains with sites of receptor tyrosine phosphorylation and are available as substrates for the activated JAKs. Following phosphorylation, the STATs form dimers through the intermolecular association of the SH2 domains with carboxyl sites of tyrosine phosphorylation. Dimerization is hypothesized to trigger dissociation from the receptor complex and translocation to the nucleus. In the nucleus, the STAT dimers bind response elements and are generally associated with the activation of gene expression.

location. One often neglected concept, predicted from this model, is that specificity in signaling is lost if the active kinases are not retained within the confines of the receptor complex.

Activation of STATs by Other, Noncytokine Receptor Families

The activation of STATs has also been reported to occur through other receptor families. The most characterized is activation of STAT1 and STAT3 by the epidermal growth factor (EGF) receptor. Importantly, the EGF receptor can directly phosphorylate STATs in vitro, although little is known regarding the mechanisms by which the STATs are recruited to the receptor and activated in vivo. There are more limited reports regarding the activation of STATs by the colony stimulating factor 1 (CSF-1) receptor tyrosine kinase and the platelet-derived growth factor (PDGF) receptor tyrosine kinase, again it is unclear whether these activate STATs directly or through JAKs. There have also been at least two reports that STATs and JAKs are activated through the angiotensin receptor, a member of the G-coupled, serpentine receptor family (Bhat et al., 1994; Marrero et al., 1995). These reports have not been substantiated, nor have any other members of this receptor family been reported to activate JAKs or STATs. In general, the consistent JAK–STAT association with the cytokine receptors contrasts dramatically from other receptor families and supports the hypothesis that the JAKs and the STATs have primarily evolved within the context of the cytokine receptor superfamily.

Biological Functions of STAT Proteins

Perhaps the most intriguing questions, and speculations, deal with the biological functions of the STATs. Do the STATs play specific functional roles in cytokine responses or do STATs play multiple roles, including growth regulation? Considerable speculation has centered on STAT1 since, in addition to the IFNs, its activation seems to occur in response to a variety of growth factors including EGF, PDGF, CSF-1, and angiotensin, as well as cytokines such as growth hormone, thrombopoietin, IL-6, and IL-10. For this reason, the phenotypes of mice lacking *Stat1* reported in this issue of *Cell* are striking in demonstrating that the sole, nonredundant function of STAT1 is to regulate a set of genes that collectively provide innate immunity. More specifically, as elegantly pointed out by the authors, none of the phenotypes are seen that would be predicted if STAT1 were involved in the other responses. Perhaps equally striking is the conclusion that a function that is so critical to existence in a less-than-perfect environment relies on a single, nonredundant transcription factor.

What about the other STATs? Three STATs will most certainly prove to be highly specific for functions related to their cognate cytokines. STAT2 is only activated by IFNα/IFNβ and is likely to be essential for IFNα/IFNβ functions. STAT4 is activated by IL-12, a cytokine that influences helper T cell differentiation to cells (Th1) that primarily produce IFN following antigen stimulation. Although it is often unwise to predict knockout phenotypes, in this case it is probably safe to assume that *Stat4*$^{-/-}$ mice will lack the ability to generate a strong Th1-type response and thus may be more susceptible to certain types of infections. STAT6 is activated by IL-4, a cytokine that influences the differentiation of Th2 helper T cells, which produce IL-4, IL-5, and IL-10 in response to antigen. IL-4 also affects B cells, in part, by activating transcription that is required for immunoglobulin class switching to IgE. Again, taking some liberty, it would appear safe to predict that *Stat6*$^{-/-}$ mice will have very specific defects in these functions. If correct, one can conclude that STAT1 and STAT2 are critical for innate immunity, while STAT4 and STAT6 play important roles in acquired immunity.

A recurring, speculative theme regarding STATs is their role in cell cycle progression and cellular transformation. Since the IFNs are not mitogenic cytokines, it is unlikely that STAT1 or STAT2 activation is directly involved in regulation of cellular proliferation. In the case of STAT6, IL-4 receptor mutants exist that fail to activate STAT6 but retain mitogenic activity, and conversely, receptor mutants exist that are nonmitogenic but retain the ability to activate STAT6. Thus, it is unlikely that STAT6 contributes to cellular proliferation. Comparable studies are not available for STAT4.

That leaves STAT3 and STAT5, which colocalize chromosomally, are expressed in most cell types, and are activated by a variety of cytokines. Each has defined functions. STAT3 induces the expression of a variety of genes that dramatically increase with tissue injury and inflammation and are therefore referred to as acute phase response genes. STAT5 regulates expression of milk proteins in the response of mammary tissue to prolactin. Indeed, an intriguing question is the whether milk proteins are induced in a variety of tissues in response to all the cytokines that activate STAT5. One would suspect not and would evoke the multifactorial nature of functional transcriptional complexes as the

basis and propose the existence of obligate cell type–specific cofactors. However, it is interesting to note that several years ago the anomalous expression of α- and β-casein expression in IL-2-dependent cytotoxic T cells was reported (Grusby et al., 1990).

Do STAT3 and STAT5 have additional tissue-specific functions, including control of cell proliferation? It has been suggested that STAT3 is the v-SRC target responsible for transformation based on the observation that it is activated in v-*src*-transformed cells (Yu et al., 1995). This observation provides no more insight into the role of STAT3 in transformation, or growth control, than any of the numerous descriptions of aberrantly phosphorylated proteins in such cells. Indeed, the mitogenic response to G-CSF is not lost with receptor mutants that no longer activate STAT3. Similarly, receptor mutants exist that dissociate STAT5 activation from JAK activation and mitogenic responses. However, all these studies are only relevant to the cell lines used and cannot provide the insights obtained with appropriate gene disruption in vivo.

Since we do not yet know the phenotype of *Stat3*$^{-/-}$ or *Stat5*$^{-/-}$ mice, the functions of the newly identified Drosophila STAT (D-STAT) become quite interesting and possibly predictive. As detailed in the reports in this issue of *Cell*, two functions are suggested from the phenotype of flies that lack D-STAT; namely, a maternal function in regulating pair rule gene expression in embryos and a zygotic function for appropriate cell numbers in the imaginal tissues in larvae. A third function, control of larval hemolymph phagocytic cell proliferation, is suggested based on the observation that a dominant negative D-STAT suppresses the proliferation of larval hemolymph cells transformed by an activated JAK. These observations demonstrate that STATs can have several unique, lineage-specific functions. Moreover, they suggest that STATs may be involved in proliferation. However, the results do not allow a distinction between regulation of genes that contribute to cell cycle progression or perhaps genes that affect differentiation and thereby cell numbers. Irrespective of this, the pheno-type of the *D-Stat* knockout ensures enthusiasm for knowledge regarding the phenotypes of *Stat3*$^{-/-}$ and *Stat5*$^{-/-}$ mice.

Future Prospects

One might anticipate exciting new information regarding the STATs in several areas. First, several intriguing questions remain regarding STAT structure. Given the nature of the proposed intermolecular SH2–phosphotyrosine interactions and the requirement for the interaction of the DNA-binding domains (try making models), the molecular structures of the STAT will be of considerable interest. The function of the conserved amino-terminal domain, which is so critical to activation, will be of interest, particularly its potential contribution to STAT dimer stabilization. In addition, studies of nuclear translocation may reveal novel mechanisms for protein transport. Finally, additional studies of the highly variable, carboxy-terminal domain and possible posttranslocation modifications may address fundamental questions in transcriptional activation or, although yet to be shown, transcriptional repression.

Ironically, one of the initial paradigms of IFN signaling, the complex containing STAT1, STAT2, and p48 seems to be unique rather than a paradigm. Specifically, p48 is a family of DNA-binding proteins that includes IRF-1, IRF-2, and ICSBP. The functions of most family members are unclear, although knockouts of IRF-1 would suggest a role in innate immunity. As yet, no additional complexes have been identified containing STATs and p48 family members.

Regarding biological functions, it can be anticipated that in the very near future the phenotypes of mice deficient in the remaining STATs will be described and will thus eliminate further speculation. However, are there additional members of the STAT family to be identified, or will the STAT family be one of the smallest transcription factor families? Considerable effort has been expended in homology screening, PCR approaches, and searching databases of expressed sequences without the identification of additional members. Thus, it would seem less likely as time goes on that additional family members will emerge. Similarly, are there additional STATs in flies?

Selected Reading

Bhat, G.J., Thekkumkara, T.J., Thomas, W.G., Conrad, K.M., and Baker, K.M. (1994). J. Biol. Chem. *269*, 31443–31449.

Copeland, N.G., Gilbert, D.J., Schindler, C., Zhong, Z., Wen, Z., Darnell, J.E., Jr., Mui, A.L.-F., Miyajima, A., Quelle, F.W., Ihle, J.N., and Jenkins, N.A. (1995). Genomics *29*, 225–228.

Darnell, J.E., Jr., Kerr, I.M., and Stark, G.M. (1994). Science *264*, 1415–1421.

David, M., Petricoin, E., III, Benjamin, C., Pine, R., Weber, M.J., and Larner, A.C. (1995). Science *269*, 1721–1723.

Durbin, J.E., Hackenmiller, R., Simon, M.C., and Levy, D.E. (1996). Cell, this issue.

Grusby, M.J., Mitchell, S.C., Nabavi, N., and Glimcher, L.H. (1990). Proc. Natl. Acad. Sci. USA *87*, 6897–6901.

Heim, M.H., Kerr, I.M., Stark, G.R., and Darnell, J.E., Jr. (1995). Science *267*, 1347–1349.

Hou, X.S., Melnick, M.B., and Perrimon, N. (1996). Cell, this issue.

Ihle, J.N. (1995). Nature *377*, 591–594.

Ihle, J.N., and Kerr, I.M. (1995). Trends Genet. *11*, 69–74.

Ihle, J.N., Witthuhn, B.A., Quelle, F.W., Yamamoto, K., and Silvennoinen, O. (1995). Annu. Rev. Immunol. *13*, 369–398.

Marrero, M.B., Schieffer, B., Paxton, W.G., Heerdt, L., Berk, D.C., Delafontaine, P., and Bernstein, K.E. (1995). Nature *375*, 247–250, 1995.

Meraz, M.A., White, J.M., Sheehan, K.C.-F., Bach, E.A., Rodig, S.J., Dighe, A.S., Kaplan, D.H., Riley, J.K., Greenlund, A.C., Campbell, D., Carver-Moore, K., DuBois, R.N., Clark, R., Aguet, M., and Schreiber, R.D. (1996). Cell, this issue.

Quelle, F.W., Wang, D., Nosaka, T., Thierfelder, W.E., Stravopodis, D., Weinstein, Y., and Ihle, J.N. (1996). Mol. Cell. Biol. in press.

Schindler, C., and Darnell, J.E., Jr. (1995). Annu. Rev. Biochem. *64*, 621–651.

Stahl, N., Farruggella, T.J., Boulton, T.G., Zhong, Z., Darnell, J.E., Jr., and Yancopoulos G.D. (1995). Science *267*, 1349–1353.

Wen, Z., Zhong, Z., and Darnell, J.E., Jr. (1995). Cell *82*, 241–250.

Yan, R., Small, S., Desplan, C., Dearolf, C.R., Darnell, J.E., Jr. (1996). Cell, this issue.

Yu, C.-L., Meyer, D.J., Campbell, G.S., Larner, A.C., Carter-Su, C., Schwartz, J., and Jove, R. (1995). Science *269*, 81–83.

Zhang, Z., Blenis, J., Li, H., Schindler, C., and Chen-Kiang, S. (1995). Science *267*, 1990–1994.